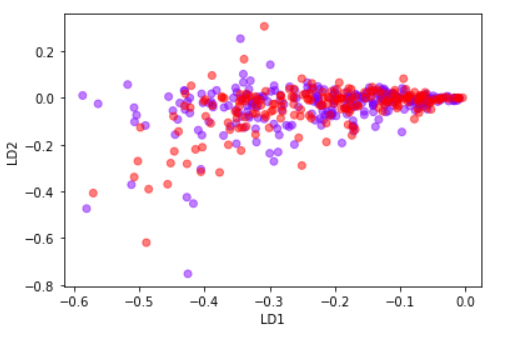
**Assignment 1 LDA and MDA**

To start off, my project is about multi label classification, however, I was not able to understand how to implement multi label classification for this assignment, so I edited the dataset a little to conform with the needs of this assignment. These changes will be mentioned.

**FLDA:**

For my implementation of FLDA I edited the dataset in a way as to make it so that there was only one genre labeled to every movie poster, that way it was no longer a multi label problem. Next, I also reduced the number of genres to only 2, Adventure and Action, these were randomly assigned to all movie posters which had genres different from these two. That way, the dataset was also more even, as both genres have, most of the time, similar number of movies posters.

For my features, I used an inbuilt function to calculate color histograms for the posters, as I wanted to differentiate them on their colours. 100 features were calculated for every poster and stored along with another variable which stored each feature vectors labels.

 After the application of LDA it was quite clear that my function did not work properly and produced a very poor, to no separation, between the classes. Hence, I was not able to calculate threshold.

**MDA:**

For my implementation of MDA I edited the dataset, once more, to have each poster only have one genre, however this time kept all the different genres that were left, instead of only keeping two, as this is a multi-class implementation. There were a total of 17 classes (19 were found by the code but two were just empty folders which were ignored during training and testing) these were Action, Adventure, Animation, Biography, Comedy, Crime, Documentary, Drama, Family, Fantasy, Horror, Musical, Mystery, Romance, Sci-Fi, Short, Thriller. They were, however, imbalanced on the number of posters each genre had.

For my features, the same as for LDA, I used an inbuilt function to calculate color histograms for the posters, as I wanted to differentiate them on their colours. 100 features were calculated for every poster and stored along with another variable which stored each feature vectors labels.

After the application of SVM classifiers, I received an accuracy of 36.65%.